

A Generalized Matrix Version of Rennie's Inequality*

John Z. Hearon**

(April 24, 1967)

The matrix version of Rennie's inequality and the finite-dimensional version of Kantorovich's inequality are obtained by considering a positive definite matrix and its inverse. Generalizations of these inequalities are obtained in which the inverse matrix is replaced by a generalized inverse with certain prescribed properties. From the generalization of the Kantorovich inequality follows a (finite-dimensional) generalization of an inequality due to Strang.

Key Words: Generalized inverse, inequality, matrix.

1. Introduction

It is the main purpose of this paper to derive an inequality which may be characterized as a generalization of Mond's matrix version [5]¹ of Rennie's inequality [7] and from which a generalization of the inequality of Kantorovich [4] follows. The inequality of Mond and some versions of the inequality of Kantorovich are derived by considering a positive definite matrix and its inverse. The derivation given here considers a positive semidefinite matrix and a generalized inverse of that matrix. The generalized inverse is required to have certain prescribed properties.

We begin with some preliminary considerations of generalized inverses of various classes. The main theorem is proved under the assumption that the generalized inverse, B , is in the weakest class and it is shown how the particular required properties of B may be weakened when B is in a stronger class.

Thus far there seems to be no treatment of the case of strict equality for Mond's matrix inequality. In the case of the Kantorovich inequality the existing discussions of strict equality are rather involved (e.g., see [3] and references therein). In any event the discussion of equality is more involved than the derivation of the inequality itself. One feature of the derivations given here is that they enable a relatively straightforward discussion of the case of strict equality for both Mond's and Kantorovich's inequalities.

2. Notation and Preliminaries

We consider matrices with complex entries and for any matrix M denote by $\rho(M)$, $N(M)$, $R(M)$ and M^* the rank, null space, range and conjugate transpose, respectively, of M . As before [2] we define $C_1(A)$ to

be the set of all matrices B such that $ABA = A$, where A is a given, arbitrary matrix, $C_2(A)$ to be the set of all matrices B such that $BeC_1(A)$ and $BAB = B$. We further define $C_3(A)$ to be the set of all matrices B such that $BeC_2(A)$ and AB is a hermitian matrix, and $C_4(A)$ as the set of all matrices B such that $BeC_3(A)$ and BA is hermitian. We call a matrix $BeC_i(A)$ a C_i -inverse of A . Other terminology has been used (see [9] and references therein and [8] and referee's suggestion and references therein). Note that a C_3 -inverse is essentially² the weak-generalized-inverse of Goldman and Zelen [1] and that the C_4 -inverse is the Moore-Penrose generalized inverse [6]. We make repeated use of the following facts: If $BeC_1(A)$, then $\rho(B) \geq \rho(A) = \rho(AB) = \rho(BA)$ and AB and BA are idempotent matrices [9]. Moreover if $BeC_1(A)$, then $BeC_2(A)$ if and only if $\rho(A) = \rho(B)$ [2, 9].

For $i = 1$ or 2 , a C_i -inverse of a hermitian matrix is not necessarily hermitian, but for every hermitian A there exists a hermitian C_i -inverse, $i = 1, 2$. In fact given $A = A^*$ and any $B_1 \in C_1(A)$ the matrix $H = (B_1^* + B_1)/2$ is in $C_1(A)$ and is hermitian [9], and the matrix HAB is in $C_2(A)$ and is hermitian [2].

In what follows we need to consider the situation $BeC_1(A)$ and $AB = BA$. Although we consider only the case $A = A^*$ and $B = B^*$, which is rather transparent, we prove the following lemma for completeness.

LEMMA 1. *If $N(A) = N(A^*)$, $BeC_1(A)$ and $E = AB = BA$, then E is an orthogonal projection and $x^*Ex = x^*x$ if and only if $x \in R(A)$.*

PROOF. Let $BeC_1(A)$, then $\rho(A) = \rho(AB) = \rho(BA)$ and if $E = AB = BA$ it follows that $N(B) \subseteq N(A) = N(E)$. From $ABA = A$, we have $R(E) = R(A)$. If $N(A) = N(A^*)$ then $N(A)$ is orthogonal to $R(A)$ and E is a hermitian, idempotent matrix. Any x can be expressed as

* An invited paper.

** Present address: Mathematical Research Branch, NIAMD, National Institutes of Health, Bethesda, Md. 20014.

¹Figures in brackets indicate the literature references at the end of this paper.

²Goldman and Zelen [1] define a weak generalized inverse B such that $BeC_2(A)$ and BA is hermitian. We follow Rohde [9] in defining $BeC_3(A)$ as in the text. Actually any statement in this paper with respect to a $BeC_3(A)$ as defined here is true for the Goldman-Zelen weak generalized inverse.

$x = u + v$, with $u \in R(A)$ and $v \in N(A)$. Then $x^*Ex = u^*u$. But $u^*u = x^*x$ if and only if $x \in R(A)$.

3. The Inequalities of Rennie, Kantorovich, and Strang

We now prove the following theorem, which is the main result of this paper.

THEOREM 1: Let A be a hermitian positive semi-definite matrix of order n and rank r . Let those roots of A which are strictly positive be ordered as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ and define $s = \lambda_1 + \lambda_r$, $p = \lambda_1 \lambda_r$. If B is any hermitian C_1 -inverse of A which commutes with A and has $v = n - r$ nonpositive roots, then for any x such that $x^*x = 1$ we have

$$x^*Ax + px^*Bx \leq s \quad (i)$$

and

$$(x^*Ax)(x^*Bx) \leq s^2/4p. \quad (ii)$$

If $\lambda_1 = \lambda_r = \lambda$, there is strict equality in (i) if and only if there is strict equality in (ii), and if and only if $Ax = \lambda x$. If $\lambda_1 > \lambda_r$, there is strict equality in (i) if and only if $x = u_1 + u_r$, where $Au_1 = \lambda_1 u_1$ and $Au_r = \lambda_r u_r$; and there is strict equality in (ii) if and only if $x = u_1 + u_r$ and $u_1^*u_1 = u_r^*u_r = 1/2$.

PROOF: Since A and B are hermitian and commute they have in common a complete orthonormal set of eigenvectors which we denote by x_i , indexed so that $Ax_i = \lambda_i x_i$, $1 \leq i \leq n$. We also have $Bx_i = x_i/\lambda_i$, $1 \leq i \leq r$. Let the v remaining roots of B be $-d_i$, $n \geq i > r$, $d_i \geq 0$. It is then clear that the matrix $P = (A - \lambda_1 I)(A - \lambda_r I)B$ is hermitian negative semidefinite. For the vectors x_i are eigenvectors of P with corresponding (nonpositive) eigenvalues $\mu_i = (\lambda_i - \lambda_1)(\lambda_i - \lambda_r)/\lambda_i$, $1 \leq i \leq r$, and $\mu_i = -pd_i$, $r < i \leq n$. From $B \in C_1(A)$ and $AB = BA = E$ we have $A = A^2B = BA^2$ and P may be written $P = A - sE + pB$. For any x , $x^*Px \leq 0$ and if $x^*x = 1$, then

$$x^*Ax + p(x^*Bx) \leq s(x^*Ex) \leq s. \quad (1)$$

Thus (i) is proved. Define $\psi = x^*Ax$ and $p(x^*Bx) = \varphi$. Then, from (1), $\psi + \varphi \leq s$ and since $\psi \geq 0$, we have

$$\psi\varphi \leq s\psi - \psi^2 \leq s^2/4 \quad (2)$$

which gives (ii).

Let e_1 denote the condition: strict equality obtains in (i); and let e_2 denote the condition: strict equality obtains in (ii).

If $\lambda_1 = \lambda_r = \lambda$ we have to prove $e_1 \Leftrightarrow e_2 \Leftrightarrow Ax = \lambda x$. We show $Ax = \lambda x$ implies both e_1 and e_2 ; then $e_1 \Rightarrow Ax = \lambda x$; then $e_2 \Rightarrow e_1$. If $Ax = \lambda x$, then it is easily verified that $\psi = \varphi = \lambda = s/2$, and we have e_1 and e_2 . Now assume e_1 , then from (1), $x^*Ex = 1$ which, by Lemma 1, implies $x \in R(A)$ and $x \neq 0$. But every $0 \neq x \in R(A)$ obeys $Ax = \lambda x$. Thus $e_1 \Rightarrow Ax = \lambda x$. Now assume e_2 , then from (2) we have $\psi = \varphi = s/2$ which gives e_1 . Thus $e_2 \Rightarrow e_1$. This completes the cases of strict equality when $\lambda_1 = \lambda_r$.

If $\lambda_1 > \lambda_r$ we have to show $e_1 \Leftrightarrow x = u_1 + u_r$, where u_1 and u_r are any eigenvectors of A associated with

the eigenvalues λ_1 and λ_r respectively. We also must show $e_2 \Leftrightarrow x = u_1 + u_r$, where u_1 and u_r are, in particular, normalized so that $u_1^*u_1 = u_r^*u_r = 1/2$. Suppose $x = u_1 + u_r$ and $u_1^*u_1 = k_1$, $u_r^*u_r = k_r$. Then

$$\psi = k_1\lambda_1 + k_r\lambda_r \quad (3)$$

$$\varphi = k_1\lambda_r + k_r\lambda_1 \quad (4)$$

$$x^*Ex = k_1 + k_r = 1 \quad (5)$$

and the sum of (3) and (4) give, in view of (5), $\psi + \varphi = s$ which is e_1 . Conversely assume e_1 , then from (1) we have $x^*Px = 0$, which requires $x \in N(P)$, and $x^*Ex = 1$, which requires $x \in R(A)$. One basis for $N(P)$ consists of the following subset of the eigenvectors x_i of A : those x_i with i such that $\lambda_i = \lambda_1$; those x_k with k such that $\lambda_k = \lambda_r$; and those x_j (if any) with j such that $\lambda_j = d_j = 0$. Thus $x \in N(P)$ can be written $x = u_1 + u_r + v$ with $v \in N(A)$. But also $x \in R(A)$, which means $v^*x = 0$ and therefore (since $v^*u_1 = v^*u_r = 0$), $v = 0$. Thus $e_1 \Leftrightarrow x = u_1 + u_r$ is proved. Now assume $x = u_1 + u_r$ and $k_1 = k_2 = 1/2$. Then (3) and (4) read $\psi = s/2$ and $\varphi = s/2$, respectively, which gives e_2 . Conversely, assume e_2 , then from (2), $\psi = \varphi = s/2$, which gives e_1 . We have just seen that e_1 implies $x = u_1 + u_r$ which in turn implies (3), (4), and (5). But when $\psi = \varphi = s/2$ any two of (3), (4), and (5) give $k_1 = k_r = 1/2$. This completes the proof of the theorem.

In case $r = n$, clearly $B = A^{-1}$ and from (i) of the theorem we recover Mond's matrix version [5] of Rennie's inequality [7]; from (ii) of the theorem we have the inequality of Kantorovich [4]. In this case we also have, from Cauchy's inequality, $(x^*Ax)(x^*A^{-1}x) \geq 1$. In case $r < n$, we have $f(x) = (x^*Ax)(x^*Bx) \geq 1$ whenever $x \in R(A)$, but this inequality is in general false. We may infringe it with the choice $x = (x_1 + x_n)/\sqrt{2}$ which gives $f(x) = (1 - \lambda_1 d_n)/4 \leq 1/4$. In fact $f(x) \geq 1$ if and only if $r = n$. For if $r = n$ we have just seen that $f(x) \geq 1$. If $f(x) \geq 1$ for all $x^*x = 1$ then $x^*Ax \neq 0$ for all $x^*x = 1$, in particular $x_i^*Ax_i \neq 0$ for all i , and A is strictly positive definite.

Preparatory to the discussion of alternative specifications of the matrix B in Theorem 1, we prove the following theorem.

THEOREM 2. Let A be a given matrix. (i) If B is a C_1 -inverse of A and commutes with A , then $E = AB$ is uniquely determined. (ii) If B is a C_2 -inverse of A and commutes with A , then B is unique and is a polynomial in A . If A is normal (in particular hermitian) then B is normal (in particular hermitian).

PROOF. (i) From $ABA = A$ and $AB = BA$, we have $BA^2 = A$, which shows that $\rho(A) = \rho(A^2)$. We further have $Ex = ABx = x$ when $x \in R(A)$, and that $Ey = AB y = BAy = 0$ when $y \in N(A)$. Since $\rho(A) = \rho(A^2)$ requires that $R(A)$ and $N(A)$ be disjoint subspaces, E is uniquely determined by A . (ii) Let B_1 and B_2 be C_2 -inverses which commute with A . From (i), $E = B_1A = B_2A$ and hence

³ If $x \in R(A)$, then $x = \sum_{i=1}^r \alpha_i x_i$. Thus

$$f(x) = \left(\sum_{i=1}^r |\alpha_i|^2 \lambda_i \right) \left(\sum_{i=1}^r |\alpha_i|^2 / \lambda_i \right) \geq 1$$

follows from Cauchy's inequality.

$B_1x = B_2x$ when $x \in R(A)$. Since [9, 2], $\rho(A) = \rho(AB_1) = \rho(AB_2) = \rho(B_1) = \rho(B_2)$, it follows from $E = AB_1 = AB_2 = B_1A$ that $N(B_1) = N(B_2) = N(A)$ and hence that $B_1y = B_2y = 0$ when $y \in N(A)$. Thus $B_1 = B_2$. From $\rho(A) = \rho(A^2)$ we know that $A = Q^{-1} \text{diag}(A_1, \theta)Q$, where $\rho(A_1) = \rho(A)$ and A_1 is nonsingular and θ is a square zero matrix. Let $p(t)$ be the polynomial such that $p(A_1) = A_1^{-1}$. Then $p(A) = Q^{-1} \text{diag}(A_1^{-1}, K)Q = G$, where $K = p(\theta)$, is a C_1 -inverse of A . For, GA is idempotent and $\rho(GA) = \rho(A)$ and this is known [2] to be necessary and sufficient for $G \in C_1(A)$. Moreover, $GAG = B$ is a polynomial in A , and [2] further $B \in C_2(A)$. If A is normal clearly B is normal. If $A = A^*$ then the coefficients of $p(A)$ are real and B is hermitian.⁴

In the proof of Theorem 2, we have seen that $B \in C_1(A)$ and $AB = BA$ together imply $\rho(A) = \rho(A^2)$. Given this rank condition on A , we have seen how to construct a polynomial in A which is a C_1 -inverse of A and, from this, a polynomial in A which is a C_2 -inverse and uniquely determined among those $B \in C_1(A)$ which commute with A . We set this forth as the following

COROLLARY: *There exists matrices $B \in C_1(A)$ which commute with A if and only if $\rho(A) = \rho(A^2)$. In this case, there exists a unique $B \in C_2(A)$ which is a polynomial in A .*

The condition in Theorem 1 that ν roots of B are nonpositive cannot be omitted if $\rho(B) > \rho(A)$. For if $x = (x_1 + x_n)/\sqrt{2}$ we have $x^*Bx = (1 - \lambda_1 d_n)/2\lambda_1$. If $d_n < 0$, B has a strictly positive root and both (i) and (ii) are violated when $-d_n > (s^2 - p)/\lambda_1 p$.

If in Theorem 1 we specify that $B \in C_2(A)$, then the condition that B be hermitian and the condition on the roots may be dropped. For, by Theorem 2, if $B \in C_2(A)$ and commutes with A then B is a polynomial in A . Further since A is hermitian, that polynomial is hermitian. But then we have $B \in C_4(A)$, since A and B are hermitian and commute and $AB = BA$ is hermitian. If we specify $B \in C_4(A)$ then the theorem holds true with no further conditions on B . For, if $B \in C_4(A)$ and $A = A^*$ it is known [6] that $B = B^*$ and $AB = BA$. If we require $B \in C_3(A)$ and $B = B^*$, then no further conditions on B are required for the validity of the theorem. For, if $B \in C_3(A)$, $A = A^*$ and $B = B^*$ then in fact $B \in C_4(A)$, since $(AB)^* = AB$ implies BA is hermitian.

If we consider the conditions of Theorem 1: (a) B is hermitian, (b) B commutes with A , and (c) B has $\nu = n - r$ nonpositive roots, then we have the following summary statement: Theorem 1 remains valid when the statements there regarding B are replaced by $B \in C_2(A)$ and (b); and when they are replaced by $B \in C_3(A)$ and (a); and when they are replaced by $B \in C_4(A)$.

It is of interest to note that if in Theorem 1 the conditions on B are replaced by the conditions that B is normal and $B \in C_3(A)$, then the theorem is valid. This follows from

LEMMA 2. *If A is normal and $B \in C_3(A)$, then A and B commute if and only if B is normal and if and only if $B \in C_4(A)$.*

⁴It is also clear that we have $B = Q^{-1} \text{diag}(A^{-1} \theta)Q$. For this matrix is clearly a C_2 -inverse which commutes with A and we have just shown that there is only one such matrix. From this it follows that A normal (in particular hermitian) implies B normal (in particular hermitian) since in these cases Q is unitary and A_1 is normal (in particular hermitian).

PROOF. Let A be normal, $\rho(A) = r$ and $Ax_i = \lambda_i x_i$, $\lambda_i \neq 0$, $1 \leq i \leq r$. Then $B \in C_3(A) \Rightarrow ABA = A \Rightarrow ABx_i = x_i$. But $B \in C_3(A) \Rightarrow AB = (AB)^*$ and thus we have $x_i^* AB = x_i^*$ or $x_i^* B = x_i^* / \lambda_i$, for $1 \leq i \leq r$. If we write $E = AB = E^*$ then since $\rho(E) = \rho(A) = \rho(B)$ we have $N(E) = N(A) = N(B)$. If B is normal then $N(B) = N(B^*)$ and we have thus proved that A and B have a complete set of orthonormal eigenvectors in common and hence they commute. Now let $AB = BA$. Then since AB is hermitian by hypothesis, BA is hermitian and $B \in C_4(A)$, further by a known theorem [6] B is normal.

It is clear from Lemma 2 that if A is hermitian and $B \in C_3(A)$ is normal, then B is hermitian, for as we have seen $B \in C_4(A)$ and this implies [6] that $B = B^*$. Thus if $B \in C_3(A)$ and is normal no other conditions on B are necessary for the validity of Theorem 1.

Finally, by use of (i) of Theorem 1 we can generalize the finite-dimensional version of a result of Strang [10] (this possibility was pointed out to the author by John W. Evans).

THEOREM 3. *Let T be an arbitrary square matrix of rank r . Let those singular values of T which are strictly positive be ordered as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ and define $s = \lambda_1 + \lambda_r$, $p = \lambda_1 \lambda_r$. If L is the C_4 -inverse of T , then, for all x , y*

$$|(y^*Tx)(y^*L^*x)| \leq \frac{s^2}{4p} (x^*x)(y^*y). \quad (6)$$

PROOF. Let $T = QA$ be the polar factorization of T , where Q is unitary and A is positive semidefinite. If B is the C_4 -inverse of A then $L = BQ^*$ is the C_4 -inverse of T . More, B is positive semidefinite and commutes with A . For it is known [6] that B is hermitian and commutes with A and hence [9], as seen in the proof of Lemma 1, that the nonzero roots of B are reciprocals of the nonzero roots of A . Thus A and B meet the conditions of Theorem 1. Let x and y be any vectors and define $z = Q^*y$. Then by the Schwartz inequality,

$$\begin{aligned} |(y^*Tx)(y^*L^*x)| &= |(z^*Ax)(z^*Bx)| \\ &\leq [(x^*Ax)(z^*Az)(z^*Bx)(z^*Bz)]^{1/2}. \end{aligned} \quad (7)$$

The singular values of T are the roots of A and (ii) of Theorem 1 can be written, for any x , as $(x^*Ax)(x^*Bx) \leq (s^2/4p)(x^*x)^2$. From this, (7) and $z^*z = y^*y$ we obtain (6).

REMARK: We cannot in Theorem 3, specify that L be a generalized inverse of T weaker than $C_4(T)$. It is a fact that if $T = QA$, Q is unitary, and $L = BQ^*$, then $L \in C_i(T)$ if and only if $B \in C_i(A)$, $1 \leq i \leq 4$. But it is not enough to have B meet the conditions of Theorem 1. For in using the Schwartz inequality to obtain (7) we additionally require that B be positive semidefinite. This requires that at least $B \in C_2(A)$ and, as we have seen, this results in $B \in C_4(A)$. For if B meets the conditions of Theorem 1 and is positive semidefinite, then the ν nonpositive roots are in fact zero, so that $\rho(A) = \rho(B)$ and $B \in C_2(A)$.

Addendum: The author's attention has been directed (by A. S. Householder) to the theorem of Diaz and

Metcalf [Bull. Am. Math. Soc. 69, 415–418 (1963), theorem 3, p. 417] which includes as a special case a Hilbert space version of Rennie's inequality [7]. They give a succinct statement of the condition for strict equality which, in the terminology of the proof of Theorem 1 of this paper, is: We have e_1 if and only if $x \in N(P)$.

4. References

- [1] A. J. Goldman and M. Zelen, Weak generalized inverses and minimum variance linear unbiased estimation, J. Res. NBS, **68B** (Math. and Math. Phys.) No. 4, 151–172 (1964).
- [2] J. Z. Hearon, Construction of EPr generalized inverses by inversion of nonsingular matrices, J. Res. NBS **71B**, (Math. and Math. Phys.) Nos. 2 & 3 (1967).
- [3] P. Henrici, Two remarks on the Kantorovich inequality, Am. Math. Monthly **68**, 904–906 (1961).
- [4] L. V. Kantorovich, Functional analysis and applied mathematics, Usp. Mat. Nauk. **3** (1948). Translation by C. D. Benster, NBS rpt. 1509 (Mar. 1952).
- [5] B. Mond, A matrix version of Rennie's generalization of Kantorovich's inequality, Proc. Am. Math. Soc. **16**, 1131 (1965).
- [6] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc. **51**, 406–418 (1955).
- [7] B. C. Rennie, An equality which includes that of Kantorovich, Am. Math. Monthly **70**, 982 (1963).
- [8] C. A. Rohde, Generalized inverses of partitioned matrices, J. Soc. Ind. Appl. Math. **13**, 1033–1035 (1965).
- [9] C. A. Rohde, Some results on generalized inverses, SIAM Rev. **8**, 201–205 (1966).
- [10] W. G. Strang, On the Kantorovich inequality, Proc. Am. Math. Soc. **11**, 468 (1960).

(Paper 71B2&3–199)